PERGAMON

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Physics and Chemistry of the Earth 28 (2003) 219–224

PHYSICS and CHEMISTRY of the EARTH

www.elsevier.com/locate/pce

Forecast of water demand in Weinan City in China using WDF-ANN model

Junguo Liu [a,*], Hubert H.G. Savenije [a], Jianxin Xu [b]

[a] *Department of Water and Environmental Resources Management, IHE, P.O. Box 3015, 2601 DA Delft, The Netherlands*

[b] *North China Institute of Water Conservancy and Hydroelectric Power, China*

Received 1 May 2002; received in revised form 1 August 2002; accepted 20 August 2002

Abstract

Domestic water use is generally the most important component of urban water consumption. In this paper, the relatively new technique of artificial neural network (ANN) is proposed to model and forecast the water demand in urban areas. Results indicate that the WDF-ANN (water demand forecast using artificial neural network) model offers an effective way to formulate domestic water demand in Weinan City in China. The model evaluation shows that the correlation coefficients are more than 90% both for the training data and the testing data.

Keywords: Water demand; Artificial neural networks; WDF-ANN model; Forecast

1. Introduction

Domestic water uses cover all uses of water by households, both within and immediately outside the confines of a residence. Domestic water demand may constitute well over half of the total municipal use in many communities (Kindler and Russell, 1984).

Because domestic water use is generally the most important component of urban water consumption, it is worth examining accurately by a proper methodology. Time extrapolation method enjoyed some popularity early in the 20th century, but is rarely used today. In multivariate models, the initial list of variables does not imply any theory of water use, so it may well omit potentially important relationships (Baumann et al., 1997). IWR-MAIN (Institute for Water Resources—Municipal And Industrial Needs) is based on observed relationships between water use and causal factors, or determinants, of urban demand for water (Baumann et al., 1997). Although IWR-MAIN model is able to accurately predict water use under a wide range of circumstances, it is expensive to be obtained and maintained. Furthermore, all the default coefficients and elasticities in this model are suitable for the USA, but they are not suitable for other places, such as for China. Also, to calibrate these coefficients and elasticities is not easy.

This paper intends to introduce a new model, WDF-ANN (water demand forecast using artificial neural network), to estimate water demand. This model combines the technology of artificial neural network and the techniques of econometrics. It does not require so many input data as in IWR-MAIN model. However, the comparison with the measured water demand and the estimated water demand by this model shows that the model is effective in estimating domestic water demand in Weinan City in China.

This paper is organized in the following way. First the background of artificial neural networks is introduced, secondly the mathematical model for water demand forecasting is developed; then the WDF-ANN model is introduced concisely. After that a case study in Weinan City is analyzed. In the last section, some conclusions are drawn based on the case study.

2. Methodology

2.1. Background of artificial neural network

An artificial neural network (ANN) is a biologically inspired computational system that relies on the collective behavior of a large number of processing elements

* Corresponding author. Tel.: +41-18235012; fax: +41-18235375.
E-mail address: liu_junguo@hotmail.com (J. Liu).

1474-7065/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S1474-7065(03)00026-3

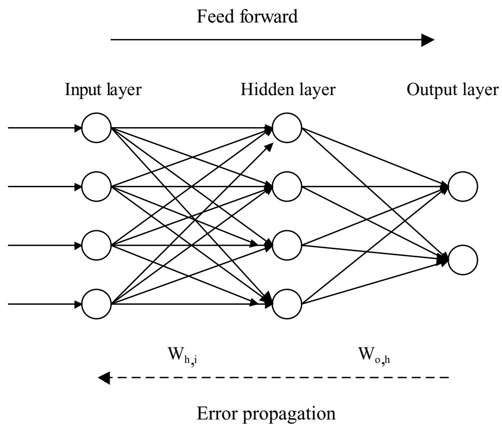

Fig. 1. Typical three-layer backpropagation neural network.

(called neurons), which are interconnected in some information-passing settings (Hassan, 2001). Lately, the ANN technique has received a great deal of attention as a tool of modeling and forecasting by many researchers and scientists. The basic idea of an ANN is that the network learns from the input data and the associated output data, which is commonly known as the generalization ability of the ANN.

The most commonly applied ANN is a backpropagation ANN as in Fig. 1, which shows a typical three-layer backpropagation neural network: an input layer consisting of four nodes, a hidden layer also consisting of four nodes, and an output layer consisting of two nodes. A numeric weight is associated with each of the inter-node connections. Weight of $W_{\mathrm{h,i}}$ represents the strength of connections of nodes between input layer and hidden layer while $W_{\mathrm{o,h}}$ represents the strength of connections of nodes between hidden and output layer.

In a backpropagation ANN, the input vector is first fed forward through the hidden layer to the output layer by using a non-linear function called activation function, which should be continuous, differentiable and bounded from above and below, then, the calculated error between the target value and the calculated output value from ANN model is back propagated through the network and the weights are updated using some training mechanism such as "generalised delta rule" (Rumelhart et al., 1986). The process of "feed forward" and "error backpropagation" is repeated until an acceptable error is attained.

The sigmoid function is usually used as the activation function because of the convenient mathematical expression of its derivate. The sigmoid function is expressed as following:

$$F(x) = \frac{1}{1 + \mathrm{e}^{-x}} \tag{1}$$

where x is the input and $F(x)$ is the output (Hassoun, 1995). It can be derived that $F'(x) = F(x)(1 - F(x))$.

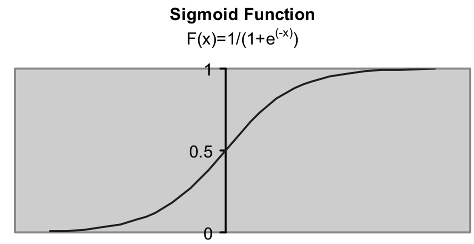

Fig. 2. Sigmoid transfer function.

This function allows simplification in deriving training algorithms. The graph of sigmoid function is as shown as in Fig. 2.

The backpropagation is essentially a gradient decent method to minimize the network error function:

$$E = \sum_{j=1}^{K} \sum_{i=1}^{N} (e_i(j) - t_i(j)) \tag{2}$$

where $e_i(j)$ and $t_i(j)$ are the estimated and targeted value, respectively. N is the number of output nodes and K is the number of training simples.

Before starting the training of an ANN, the weights are initially randomized. Based on the error propagation, the weights are adjusted based on equation:

$$\Delta W_{ij}(n) = \alpha \frac{\partial E}{\partial W_{ij}} + \eta \Delta W_{ij}(n-1) \tag{3}$$

where $\Delta W_{ij}(n)$ and $\Delta W_{ij}(n-1)$ are the weights increment between node i and j during the adjacent iteration n and $(n-1)$. α and η are learning rate and momentum factor. Generally speaking, a small learning rate can ensure the reduction of the error function but may slow the convergent process, while a large learning rate can speed the learning process but may cause divergence (Qi et al., 1997). The use of momentum is able to stabilize the solution and help the network to learn fast since part of previous weight change is applied to the current weight change. Careful selection and appropriate adjustments of the learning rate are necessary to successfully train the backpropagation neural network.

2.2. Mathematical model for water demand forecast

In this paper, a three-layer ANN was designed to process an input vector consisting of water price (P), house income (I), and household size (S) in order to generate an output vector, water demand (D, water consumption per capita per day, with a unit of m^3/cap/day).

Price and household income are recognized to be the usual variables that should be considered in any demand function. The water price is very important for water demand management to achieve effective and sustainable use. The household income determines the willingness to pay for the water users. For the same household income, the larger the family is, the more water the household will consume. But given the same income per capita, the larger the family is, the smaller the total water demand is. This is because in a larger family, it often means that there are more old or young people who cannot consume as much water as the young for bathing, food and toilet uses etc.

According to Goodman (1984), the economic value of a product or service from a water resource project is correctly estimated as the amount users are willing to pay for it. The willingness to pay of water users is a function of the quantity that users consume and their ability to pay. The following conclusions could be drawn when water demand is to be forecasted:

- generally speaking, water demand (the quantity that users consume) will decrease as *water price* increases;
- the rich have a higher ability to pay than the poor.

Standardization is required to normalize the input vector and the output vector in the model development. For the input vector, due to the different units (for example, the unit of water price is Y/m^3, the unit of household size is capita, however the unit of household income is Y/yr. The currency unit in this paper is Chinese Yuan (Y). The exchange rate was 1 US$ = 8.279 Y) and different magnitude (the value of water price is about 1.2, the household size is about 3, but the household income is about 4000 in Weinan City), it is required to normalize all the input data in the range of 0–1. For the output vector (water demand, with a unit of m^3/cap/day), based on the characteristic of sigmoild transfer function, the output value is between 0 and 1. But the values of water demand are not always in this range. So standardization is also required for the output vector.

The model of water demand forecast using artificial neural network is expressed as follows:

$$D = F(W_2F(W_1R + \theta_h I) + \theta_o I) \tag{4}$$

where D is the standardized water demand (output vector); R is the standardized input vector, which consisting of water price (P), house income (I), and household size (S); $F(x)$ is the sigmoid function; W_1 is the weight vector between input layer and the hidden layer; W_2 is the weight vector between hidden layer and the output layer; θ_h is the threshold vector of hidden layer; θ_o is the threshold vector of output layer; I, unit matrix, $I = [1, 1, \ldots 1]$.

2.3. WDF-ANN model

WDF-ANN (water demand forecast using artificial neural network) is a PC-based software tool used to make projection for domestic water demand. This model was developed to forecast water demand based on backpropragation algorithm of ANN technology. In backpropragation, the most important process is to calibrate the weights and the threshold. This process is implemented with the software package Fortran PowerStation. For the processes of verification and forecast, they are implemented with the software package Delphi 6.0. Also the interface of this model is designed with Delphi 6.0.

WDF-ANN is not a sophisticated program. It has a friendly interface as shown in Fig. 3. It can be operated by anyone with basic skills in the operation of computers, in the use of Microsoft Windows.

3. Case study

3.1. Water demand in Weinan City

Weinan City is situated in the east of Guanzhong Plain of Shaanxi Province in China. It has a total area of about 126 km^2 with a population of approximately 230,000 people, which means a population density of 1830 people/km^2. Among the whole people, 67% of them are living in urban area.

Weinan City has the lowest water resources in Shaanxi Province. Water resources per capita are 300 m^3/yr, which only makes up 1/7 of China's average and 1/28 of the world's average.

3.2. Data collection and model calibration

To train a neural network, it is necessary to have a large number of datasets of high quality. All the data derived from Weinan Housing Office in Weinan City were employed in this paper. The base year for use in creating WDF-ANN model is chosen as the year of 1998 because of the abundant socio-economic data availability in this year. Data collected represents approximately 53% of total water use in the city and the collection was accomplished with a field survey in order to obtain the following required input of WDF-ANN:

- quantity of water consumed per month (m^3/connection/month),
- water and wastewater prices in effect (Y/m^3),
- household size (capita),
- household income (Y/yr).

Fig. 4 shows the relation between water demand (domestic) and household size of different conditions of per capita income. Based on the graph, the following conclusions could be drawn for domestic water consumption in Weinan City.

- Generally speaking, the larger per capita income will lead to larger water demand.

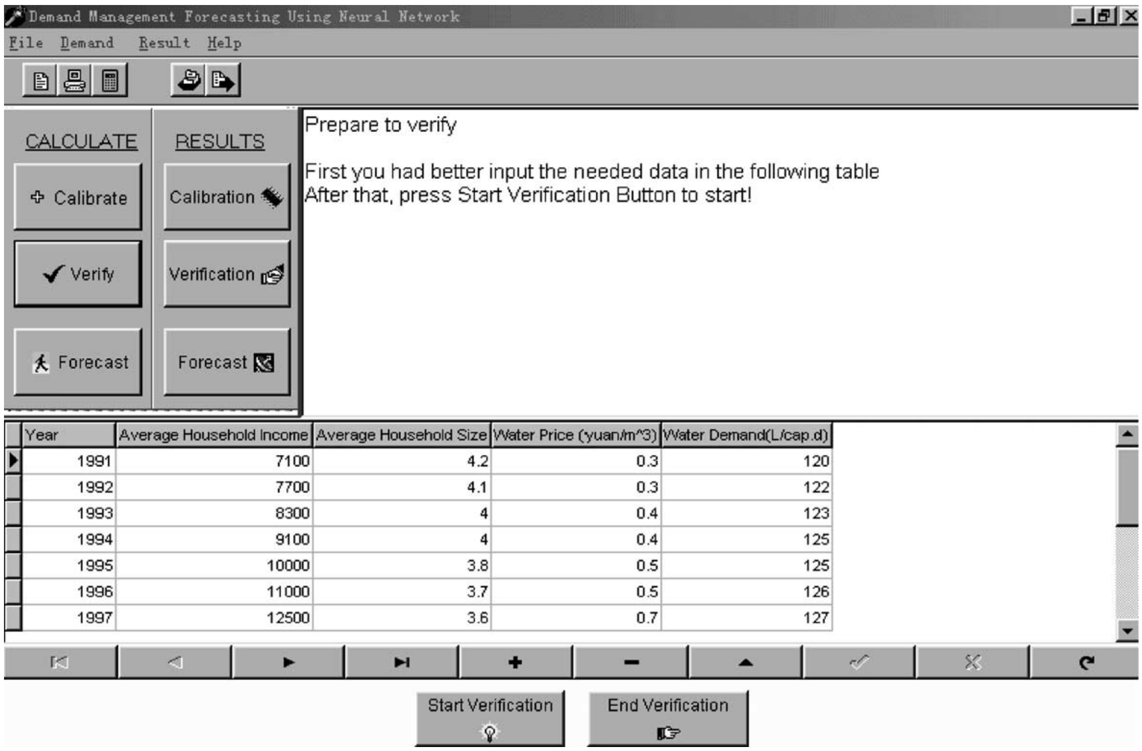

Fig. 3. Interface of WDF-ANN.

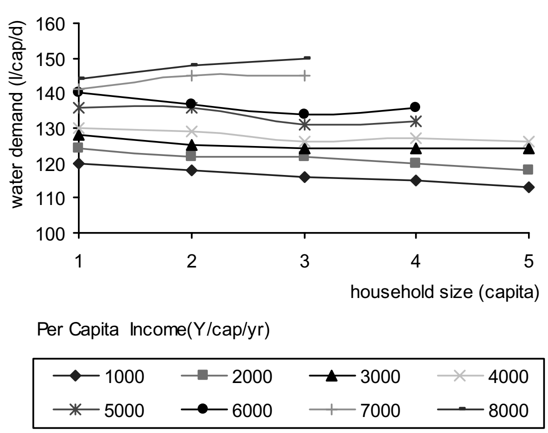

Fig. 4. Water demand–household size relation under different conditions of average income per capita.

- As to the same per capita income, normally the larger the household is, the smaller the water demand is. This is because that larger family often means that more old people or more children, who cannot consume as much water as the young people.
- For the rich families with a high per capita income, such as 7000 or 8000 Y/cap/yr, water demand will increase a little as the increase of the household size. Based on the analysis by Weinan Housing Office, around 60% high-income single families (household size is 1) prefer to have dinners in the restaurants instead of at home.
- Seldom large family could be found in the high-income family.

Fig. 5 shows the domestic water demand and household income relation for different household size. The following trends can be concluded:

- For all household sizes, higher household income results in the rise of water demand. But for the smaller household, the increase of water demand is steeper with the increase of household income.
- Water demand is more elastic for the poor than for the rich.

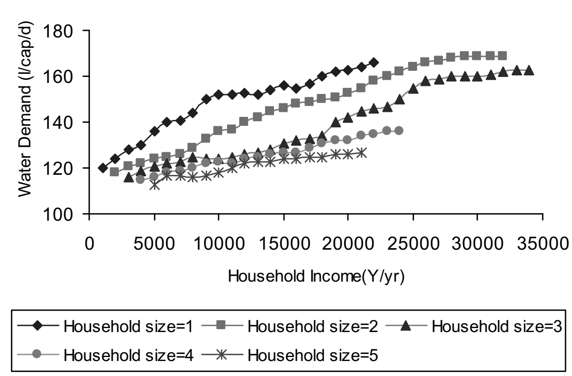

Fig. 5. Domestic water demand–household income relation under different household.

- For the median family (say family size is 3–5 capita and household income is 10,000–25,000 Y/yr), the water demand does not change a lot with the increase of household income or household size.

Based on the investigated data, the WDF-ANN model was run to calibrate the weights and threshold values. The error limit was chosen as 4×10^{-4}. The results are introduces as follows:

Loop times: 70,322 times

The weighted value and threshold as follows
WH(J,I)—weights from input node I to hidden node J
WO(K,J)—weights from hidden node J to output node K
SETAH(J)—the threshold in hidden layer
SETAO(K)—the threshold in output layer
WH(1,1) = 1.968534
WH(2,1) = 7.489122E−01
WH(3,1) = 1.367978
SETAH(1) = 1.619887
WH(1,2) = −1.996743E−01
WH(2,2) = 7.686946E−01
WH(3,2) = 1.577518
SETAH(2) = 1.734083E−01
WH(1,3) = 1.217903
WH(2,3) = 2.543249
WH(3,3) = −8.280971E−01
SETAH(3) = 3.549103
WH(1,4) = −19.029440
WH(2,4) = −2.246556
WH(3,4) = −3.761171
SETAH(4) = −3.093015
WO(1,1) = 2.251125
WO(1,2) = −7.567549E−01
WO(1,3) = 6.602859
WO(1,4) = −9.399937
SETAO(1) = 1.619977

3.3. Verification and forecast of WDF-ANN model

The next step was to use the calibrated model to estimate water demand in other years in Weinan City, as a test of model effectiveness. The average household income, average household size, water price and water demand could be obtained from 1991 to 2000. The result of verification is as shown as in Fig. 6. As a rule of thumb, differences below 3% indicate good performance; differences exceeding 5% usually mean further calibration is needed. The maximum of difference in this verification is only 2.3% in year 1991, which means the calibration is good enough to be employed to estimate water demand.

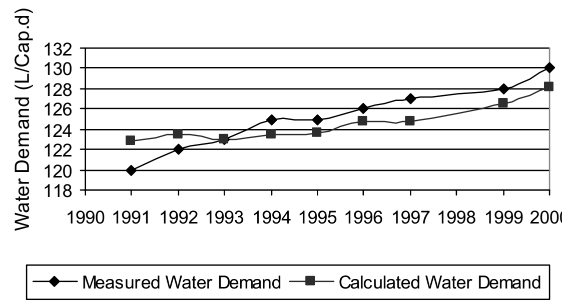

Fig. 6. Water demand comparison between the measured value and calculated value using WDF-ANN model.

It was estimated that the average household income would reach to 18,000 Y; the household size would decrease to 3.2. If the average water price were 0.8 Y/m^3 in 2005, water demand would be 138.5 l/cap/day; if the average water price were 1.0, 1.2 and 1.5, water demand would be 135.5, 133.3 and 131.7 l/cap/day.

3.4. Model evaluation

In this study, prediction error is evaluated in two aspects: the error assessment for the training data and for the testing data. The purpose of the error assessment for the training data is to evaluate ability of the model to recognize and reproduce the training data set. The purpose of the error assessment for the testing data is to check the ability of the model to predict the water demand with the calibrated coefficients. The main statistical tools are the root-mean-square error and R^2. The root-mean-square error is a statistical measure of the differences between the actual output values and the output values predicted by the network model. It is a mean over the whole data set of the square of the actual value minus the predicted value. The errors are squared to cancel the effect of the positive and negative values of the differences.

Table 1 shows the results of the evaluation of WDF-ANN model. These parameters represent how well the WDF-ANN model fits the training and testing data. It was found that the correlation coefficients both for the training data and the testing data are more than 0.9. It is large enough to indicate that the WDF-ANN model is accurate for predict water demand.

Table 1
The statistical data for WDF-ANN model

	Training data	Testing data
Data number	30	9
R^2	0.981	0.918
Correlation coefficient	0.99	0.919

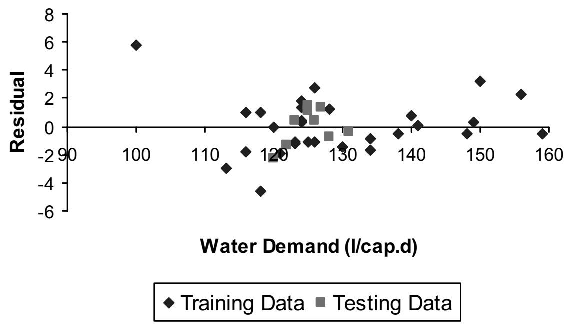

Fig. 7. The residual versus water demand.

In order to check the goodness of fit of WDF-ANN model, a residual analysis was performed. The plot of the residual versus water demand is shown in Fig. 7. From it, it can be found that both for training data and for testing data, the mean of the residuals is approximately 0. In both the two series, no special trend occurs for the distribution pattern other than the random scatter pattern. Although the density of the residuals is different in some regions, it is due to the different number of sampling data available in that region instead of bias.

Over all, WDF-ANN model is accurate enough to be employed to predict water demand in urban areas in Weinan City.

4. Conclusion

Although not much literature could be found about the ANN technique for water demand forecasting, the ANN modeling approach appears promising to predict water demand in urban areas. It is a fast and flexible way to incorporate econometric techniques into ANN model. The above analysis shows that the method of artificial neural network offers the possibility to estimate water demand for domestic use in Weinan City in China. Furthermore, it shows that the WDF-ANN model offers an effective way of accounting for domestic water demand. From the case study, the conclusion can be drawn that the results between the measured and calculated water demand by WDF-ANN model are compatible to formulate the domestic water demand in Weinan City. Application of this model in other cities is promoted to further assess the compatibility of this model.

It is always believed that any efforts are not perfect and certainly there is room for further argument and improvement. The limitation of this study is argued as follows. First, other methods, such as time extrapolation method, multivariate model method etc, are required to employ for comparison purpose. Second, it might be possible to develop different types of ANN models for further study. In this paper, a simple ANN model consisting of only one hidden layer was developed. However, complex ANN models consisting of two or more hidden layers are possible to optimize architecture of the ANN to capture the relationship among the various input and output variables. Finally, it is realized that the method developed in this study need forecasted information such as the population, income and water price. How successful the water demand model developed in this study will perform when presented with forecasted information still remains to be seen. It is hoped that the future research would concentrate in the above argument to overcome the limitation of the present research.

References

Baumann, D.D., Boland, J.J., Hanemann, W.M., 1997. Urban water demand management and planning. McGraw-Hill, New York.

Goodman, A.S., 1984. Principles of Water Resources Planning. Prentice-Hall, Englewood Cliffs, NJ.

Hassan, A.E., 2001. Prediction of plume migration in heterogeneous media using artificial neural networks. Water Resources Research 37 (3), 605–623.

Hassoun, M.H., 1995. Fundamentals of Artificial Neural Networks. The MIT Press, Cambridge.

Kindler, J., Russell, C.S., 1984. Modeling water demands. Academic Press, London.

Qi, G.Z., Yang, J.C.S., Amini, F., 1997. Neural network for identification and control of civil engineering structures. In: Kartam, N. et al. (Eds.), Artificial Neural Networks for Civil Engineers: Fundamentals and Applications. ASCE, pp. 92–123.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning internal representations by error back propagation. In: Rumelhart, D.E., McClelland, J.L. (Eds.), Foundations. In: Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1. The MIT Press. (Chapter 8).

Junguo Liu is an associate expert in CapNet (International Capacity Building Network for Integrated Water Resources Management), The Netherlands. He finished his MSc study from IHE-Delft in April 2002 with the major of Water and Environmental Resources Management. His research interests focus on hydrology and water resources, hydroinformatics, GIS and remote sensing application in water sectors, environmental science and climate change study. Email: liu_junguo@hotmail.com

Hubert H.G. Savenije is professor of water resources management at IHE-Delft, The Netherlands. He has worked for many years in the developing world on issues related to water policy, management of international rivers, flood management, water resources planning, and hydrology. He lectures at IHE, Delft University of Technology, Vrije University of Amsterdam and the University of Zimbabwe. Email: has@ihe.nl

Jianxin Xu is professor of water resources management at North China Institute of Water Conservancy and Hydroelectric Power, Zhengzhou City of China. He has worked for many years on issues related to water-saving irrigation, water resources assessment and decision support system development for water resources management. Email: xujianxin1954@163.com